

MULTIPLE TIME SCALES IS WELL NAMED

JOHN GIBBON

COLUMBIA UNIVERSITY

Staddon and Higa's article is a critique of scalar expectancy theory, and a proposed alternative, multiple time scales. The critique is generally flawed, both factually and logically. The alternative is bewildering in its flexibility, opaque in its quantitative description, and never addressed to real data.

Key words: timing, scalar expectancy theory

The article by Staddon and Higa is first a critique of, or more properly a diatribe on, scalar expectancy theory (SET), and second a proposed alternative. The paper is not satisfactory in either section, for very different reasons. The criticisms of SET are often ill taken and occasionally bizarre. On the other hand, the multiple-time-scales (MTS) alternative proposed next is a bit too aptly named. Staddon and Higa mix and match time scales to qualitative descriptions of effects in the timing literature with a bewildering variety. Power scales, log scales, summed exponential scales (MTS), and even linear scales (see below, time-left explanation) are used almost at will.

When the original version of the article was submitted to *JEAB*, the editor requested a review from me, and I submitted a signed review. The paper was subsequently revised, resubmitted, and accepted to be published with commentary. My reading of the revision is that not much has changed (with the exception of Figure 6, discussed below). I am therefore providing a commentary which is essentially portions of my original review, with discussion of minor points deleted and edited for clarity.

Poisson Variance

The authors begin with a central theme, namely that the pacemaker-accumulator idea in SET is difficult to reconcile with Weber's law. Indeed, the description implies that SET is continuously scrambling to get around this difficulty by adding parameters to the ac-

count. A central feature of a pacemaker accumulator is Poisson variance, which does not accord with the scalar property: proportional rescaling of timing distributions. Actually, from the outset the information-processing account of SET was designed with just the scalar property in mind. The idea was that a system of this kind (a) must account for the scalar property and (b) may do so in more than one way. The clock, memory, and decision process stages identified three potential sources of scalar variance, all of them with the critical property that random variation in the system be multiplicative. The three sources so identified were pacemaker rate variation (by the way, it makes no difference whether one assumes trial-to-trial variation or within-trial variation; cf. Gibbon, 1992), memory translation multipliers in storage and retrieval, and threshold variation. The authors have, I think, misunderstood the way in which the notion of a Poisson pacemaker (in both early and late versions of SET) was conceived. Gibbon (1992) showed that even small sources of multiplicative variance do indeed render Poisson variability negligible. The assumption of a Poisson pacemaker was a convenient one that is physiologically plausible, because there are many neural systems with Poisson variability. The key features of the theory did not rely on the Poisson pacemaker idea but rather on a mechanism that integrates activity over time with multiplicative variance.

Logarithmic Subjective Time Scale

Staddon and Higa then go on to posit a logarithmic perceptual subjective time scale. They argue that the temporal bisection result with indifference at the geometric mean is a natural and straightforward instantiation of equal subjective distance from the two anchor

This research was supported by NIMH Grant MH41649-12.

Address correspondence to John Gibbon, Biopsychology, Unit 50, Columbia University, 722 W. 168th St., New York, New York 10032 (E-mail: jg34@columbia.edu).

points, short and long. As Staddon and Higa state, the geometric mean finding drops out of the log time scale rather simply. As long as variance is constant and symmetric on the log scale, the midpoint between two remembered values should be at the geometric mean. In fact this was the motivation for examining the subjective time scale in the time-left experiments (see below). For bisection, given the linear scale that SET assumes, the psychophysical function is obtained not by simple differences on the subjective scale but rather by similarities on this scale, where similarity is defined, like other discriminative functions in SET, as a ratio, in this case of the probe to the referents. Staddon and Higa have not described the SET analysis of bisection accurately, in my view. Gibbon (1981) analyzed logarithmic and linear scales using a difference rule for the logarithmic scale and the ratio rule for the linear scale. Although it is obviously true that differences on a log scale reduce to ratios on a linear scale, it is not true that the form of the psychometric function is well fit by the logarithmic scale. Indeed, it is shown that symmetry of remembered distributions of the standards on the log scale results in a poor fit to the psychometric functions, albeit with an appropriate indifference point at the geometric mean.

There are other data that deal with the form of the psychophysical function as well, which again make for difficulties with the log scale. Church and Gibbon (1982) examined what form temporal generalization functions ought to take when animals discriminate whether a probe stimulus is the same or different from a reinforced standard. The scalar property is again found, but with near symmetry on the real time axis, which is not to be expected from a subjective log scale. Rather a negative skew, which is never found in real data, is predicted for the log scale.

Time-Left Experiments

The time-left experiments (Gibbon & Church, 1981) were motivated precisely by the geometric-mean finding for bisection. But Staddon and Higa do not seem to have grasped the fundamental features of these experiments, particularly Experiment 2. Subjects choosing between an elapsing alternative and another, the standard, which is fixed (beginning "right now"), should show choic-

es reflecting proximity on the log scale to food, which is much shorter in the middle of the interval for the time-left side than it is for the standard. But importantly, if a log unit is added to both sides by doubling the elapsing and standard interval, as was done in the second experiment, then a discrimination based on proximity to food on the log scale predicts *no change* in the point of indifference between the two. And this of course is ruled out by the data from Experiment 2.

Even with omission of Experiment 2, Staddon and Higa evidently find the log scale not tenable for the time-left experiments, and perform a bizarre contortion to obtain the desired result. They argue that subjects perceive time in a logarithmic fashion but then take *inverses* (i.e., antilogs) and then calculate immediacies (expectancies) in real time! That is, subjects perceive time logarithmically but have the good sense to ignore their perceptions (or the good sense to pick the antilog transform) when faced with a choice between delays to food. They use real time, as though the log scale associated with perception of time were irrelevant. SET does not need such convoluted reasoning to arrive at the appropriate result, because the point of indifference should scale linearly with real time provided that time is perceived linearly with real time.

Staddon and Higa claim a kind of "perceptual constancy mechanism" to account for the use of linear time. They argue "for a separation between the animal's capacity to assess reinforcement rates and its capacity to use a decaying memory trace as a stimulus" (p. 222). The example espoused to justify such a convoluted interpretation is that there is a smaller two-point threshold on the hand than on the back "but we do not feel that our hands are larger than our back" (p. 223). This kind of reasoning from just noticeable differences is exactly the sort that led Fechner to the log scale. Do the authors wish to challenge his inference? In any case, this example makes no sense applied to the time scale. The reason for judging that the hands are smaller than the back has nothing to do with the two-point threshold. Visual sensory data are undoubtedly used in determining size, but for temporal judgments there is no alternative receptor. What are the alternative data used when judging time durations?

The authors go on to argue for a log time scale on other grounds (p. 224). However, I confess to be completely lost on Equations 8 and 9. How is it that “internal effects” are related to time differentials by these equations? Multiplying through by t or b , they seem to be saying that a small change in t (dt) is equal to the standard deviation of t times the “internal effects” or differential on z , $dt = \sigma_t dz$. But what is z ? And where is the variance in this system? Although it is certainly true that Equations 10 and 11 follow by integration, the justification for Equations 8 and 9 remains mysterious. This is but one of several examples of what might be called mini models in this paper, none of which are developed in sufficient detail for us to evaluate them. Moreover, it is not always clear which mini model is being espoused, because power, log, and MTS are frequently given equal weight, and are introduced where convenient for one or another purpose. It is also unclear how the mini models relate to the presumably central, more developed MTS model. Sometimes, as in Figure 1, they are touted as equivalent, but of course in many other respects, such as slopes, they are clearly not.

Start/Stop/Spread Correlations in the Peak Procedure

Staddon and Higa insinuate that positive start-stop correlations (as opposed to negative start-spread correlations) are a difficulty to be overcome in SET. In fact this analysis was designed to attempt to isolate the relative contributions of memory and decision variance in this procedure. Far from being unexpected, the analysis showed that both sources of variance are present.

MTS, in contrast, appears to have only a start threshold. How would such a system handle the results of the peak procedure? If a stop threshold is to be permitted, how does the animal “know” where to place it, or even come to “expect” reinforcement at a particular time, and so detect its omission?

MTS Memory Timing Model and the Variance Problem

Figure 6 in the current article is very different from that in the original. The original Weber fraction was shown to decrease about 55% over an eightfold range in interfood in-

terval (a Poisson system would show a Weber fraction decrease of about 65% over this range). The description in the original, however, closely matched that in the current version (pp. 232–233). The text appears to describe the Weber fraction as resulting from the decreasing slope of the memory trace.

The new Figure 6 shows rising and falling functions. The degree of change near the origin is small for values of $\lambda \geq 1.5$. However, small on what scale? Ordinate values are omitted from the new Figure 6. Indeed, it would be useful to have more than simply ordinate values in the new Figure 6; it would be much more compelling to see actual Weber fractions from real data obtained over these ranges.

The problem with where variance arises in this system is endemic to all of the discussion. Staddon, at a recent meeting (personal communication), described the Weber fraction as obtained from a simulation of MTS with threshold variance. That is, a given threshold on the memory decay trace functions with constant variability would induce the curves shown in Figure 6. Is threshold variance then the source of variability producing the Weber fraction? The description in the text does not make this clear at all. Indeed, there is no talk of variance throughout this paper, and the quantitative mechanisms underlying variance, or even those underlying the scales, are generally opaque (MTS is never described with closed forms). And if the lack of detail on variance mechanisms is a problem for me, I suspect I am not alone.

It is almost as though animals are perfectly accurate but somehow the slope of the subjective time functions induces a Weber fraction without any variability (at least in Equations 19 and 20). If threshold variance is indeed what induces the Weber fraction here, then that description needs to be made explicit. Threshold variance is one of the sources of scalar variance in SET, but it is always explicitly so. Here it is not clear (a) whether the Weber fraction increases or decreases or (b) whether the functions in Figure 6 are a result of a simulation with true variability or are the result of some calculations on slopes, as the text suggests.

Even on a qualitative level, it seems that the MTS decay functions, which can approach arbitrarily close to zero over a rather short

range (see Gallistel's commentary), must pose processing problems for real live subjects that have internal processing noise; for example, constant threshold variance must avoid negative threshold values as the MTS subjective scale approaches zero. But truncating the threshold distribution at zero means that it is no longer constant, and is less so the closer the decay function is to zero.

Deterministic accounts are in principle error free, but a major thrust of psychophysics for many years has been to understand sources of variability and error. We need to know more about how errors are produced in this system before an evaluation can be made. On

its face, MTS appears to avoid the variance problem by simply remaining silent.

REFERENCES

- Church, R. M., & Gibbon, J. (1982). Temporal generalization. *Journal of Experimental Psychology: Animal Behavior Processes*, 8, 165–186.
- Gibbon, J. (1981). On the form and location of the psychometric bisection function for time. *Journal of Mathematical Psychology*, 24, 58–87.
- Gibbon, J. (1992). Ubiquity of scalar timing with a Poisson clock. *Journal of Mathematical Psychology*, 36, 283–293.
- Gibbon, J., & Church, R. M. (1981). Time-left: Linear vs. Logarithmic subjective time. *Journal of Experimental Psychology: Animal Behavior Processes*, 7, 87–108.

MODELING MODELING

PETER R. KILLEEN

ARIZONA STATE UNIVERSITY



Models are tools; they need to fit both the hand and the task. Presence or absence of a feature such as a pacemaker or a cascade is not in itself good. Or bad. Criteria for model evaluation involve benefit-cost ratios, with the numerator a function of the range of phenomena explained, goodness of fit, consistency with other nearby models, and intangibles such as beauty. The denominator is a function of complexity, the number of phenomena that must be ignored, and the effort necessary to incorporate the model into one's parlance. Neither part of the ratio can yet be evaluated for MTS, whose authors provide some cogent challenges to SET.

Key words: models, pacemakers, theories

If you think models are about the truth, or that there is a best timing model, then you are in trouble. There is no best model, any more than there is a best car model or swimsuit model, even though each of us may have our favorites. It all depends on what you want to do with the model. Nor are models theoreticians' guns of domination, any more than data are empiricists' bullets of assault. War

games can be fun, however, especially when you do not have entangling alliances with the principles, and can just watch them swat it out. Will the grapplers or the punchers win this year? It happens to empiricists too, failing to replicate and sniffing about controls, but somehow it is more fun when it is the guys in suits, the guys who prefer *ln* to *log*, the guys who try to explain *your* data to *you*, are going at it.

And it is a good thing for them to do, too. Do *you* want to attempt to tell a Gibbon that his integral is improper? Or a Staddon that he might have one too many layers in his

This work was supported by NSF Grant IBN 9408022 and NIMH Grant K05 MH01293.

Address correspondence to Peter Killeen, Department of Psychology, Box 1104, Arizona State University, Tempe, Arizona 85287-1104 (E-mail: Killeen@asu.edu).